

# Structure Preserving H-infinity Optimal PI Control

Anders Rantzer\* Carolina Lidström\*\* Richard Pates\*\*\*

\* Lund University, Sweden (e-mail: rantzer@control.lth.se).

\*\* Lund University, Sweden (e-mail: carolina.lidstrom@control.lth.se).

\*\*\* Lund University, Sweden (e-mail: richard.pates@control.lth.se).

**Abstract:** A multi-variable PI (proportional integrating) controller is proved to be optimal for an important class of control problems where performance is specified in terms of frequency weighted H-infinity norms. The problem class includes networked systems with a subsystem in each node and control action along each edge. For such systems, the optimal PI controller is decentralized in the sense that control action along a given network edge is entirely determined by states at nodes connected by that edge.

*Keywords:* Distributed control, decentralized control, linear systems, robust control

## 1. INTRODUCTION

Classical theory for multi-variable control synthesis suffers from a severe lack of scalability. Not only does the computational cost for Riccati equations and LMIs grow rapidly with the state dimension, but also implementation of the resulting controllers becomes unmanageable in large networks due to requirements for communication, computation and memory. Because of this situation, considerable research efforts have recently been devoted to development of scalable and structure preserving methods for design and implementation of networked control systems building on early contributions by Bamieh et al. (2002), D’Andrea and Dullerud (2003) and Rotkowitz and Lall (2002).

In practice, most scalable control architectures are built on layering. For example, control systems in the process industry are often organized in a hierarchical manner, where scalar PID controllers are used at the lowest level and the reference values of these controllers are computed on a slower time scale by centralized optimization algorithms. This approach often works well provided that the scalar loops are reasonably decoupled and that the coordination dynamics are comparatively slow. Other important applications where decentralized and layered control architectures have been successfully applied are power systems and Internet traffic control.

A scalable approach to control synthesis has recently been developed based on the notion of positive systems (Rantzer (2015)) and the nonlinear counterpart monotone systems (Dirr et al. (2015)). Such systems are characterized by existence of Lyapunov functions and other performance certificates whose complexity grows only linearly with the systems size. Interestingly, the maximal gain is always attained at zero frequency. Restricting attention to closed loop positive systems, distributed static controllers were optimized subject to  $H_\infty$  performance in Tanaka and Langbort (2011), while  $\ell_1$  performance was considered in Briat (2013).

The powerful synthesis methods for positive systems have raised an important question: How restrictive is a demand

for closed loop positivity in  $H_\infty$  optimal control? It was therefore a remarkable step forward when Lidström and Rantzer (2016) showed that for a large class of networked control systems with “diffusive” dynamics (i.e. symmetric state matrix), it is not restrictive at all. Instead, controllers defined by a simple closed form expression achieve the same level of performance as centralized controllers derived using Riccati equations or LMIs. In particular, the following problem was considered:

*Given a graph  $(\mathcal{V}, \mathcal{E})$  and the system*

$$\dot{x}_i = a_i x_i + \sum_{(i,j) \in \mathcal{E}} (u_{ij} - u_{ji}) + w_i \quad i \in \mathcal{V}, \quad (1)$$

*find a control law of the form  $u = Kx$  that minimizes the  $H_\infty$  norm of the transfer function from the disturbance  $w$  to the controlled output  $(x, u)$ .*

Lidström and Rantzer (2016) proved that when  $a_i < 0$  an optimal control law is given by

$$u_{ij} = x_i/a_i - x_j/a_j \quad (2)$$

and the closed loop from  $w$  to  $x$  is a positive system. This control law is trivial to compute and decentralized in the sense that control action on the edge  $(i, j)$  is entirely determined by the states in node  $i$  and node  $j$ .

The expressions above define a rare but important class of systems where decentralized controllers are known to achieve the same  $H_\infty$  performance as the best centralized ones. Still, the setting is insufficient for many practical applications. In particular, as all proportional controllers, the control law (2) is unable to remove static errors in presence of constant disturbances. The purpose of this paper is to get rid of this deficiency, by modifying the performance criterion to optimize dynamic controllers with integral action.

The structure of the paper is as follows: After introduction of some basic notation in section 2, we prove the main result in section 3, consider some basic applications in section 4 and 5, before making conclusions in section 6. A well known matrix optimization result is included for completeness as a short appendix.

## 2. NOTATION

Let  $\mathbf{RL}_\infty$  be the set of proper (bounded at infinity) rational functions with real coefficients. The set of  $m \times n$  matrices with elements in  $\mathbf{RL}_\infty$  is denoted  $\mathbf{RL}_\infty^{m \times n}$ . Given  $P \in \mathbf{RL}_\infty^{k \times l}$ , we say that  $K \in \mathbf{RL}_\infty^{l \times k}$  is *stabilizing*  $P$  provided that

$$\begin{bmatrix} I \\ K \end{bmatrix} [I + PK]^{-1} [I \ P] \in \mathbf{RH}_\infty^{(k+l) \times (k+l)}$$

has no poles in the closed right half plane. Furthermore  $\|P\|_\infty = \sup_{\text{Re } s \geq 0} \|P(s)\|$ . For a matrix  $M \in \mathbb{R}^{n \times m}$ , we denote the pseudo-inverse by  $M^\dagger$  and the spectral norm by  $\|M\|$ . For a square symmetric matrix  $M$  the notation  $M \succ 0$  means that  $M$  is positive definite, while  $M \prec 0$  means that  $M$  is negative definite.

## 3. MAIN RESULT

*Theorem 1.* Let  $P(s) = (sI - A)^{-1}B$  with  $A$  symmetric negative definite. Assume that  $\tau \geq \sqrt{\|B^T A^{-4} B\|}$ . Then the problem

$$\begin{aligned} &\text{Minimize} \quad \|(I + KP)^{-1} K\|_\infty \\ &\text{subject to} \quad \|\frac{1}{s}P(I + KP)^{-1}\|_\infty \leq \tau \end{aligned}$$

over stabilizing  $K$ , is solved by

$$\hat{K}(s) = k \left( B^T A^{-2} - \frac{1}{s} B^T A^{-1} \right)$$

where  $k = \|(A^{-1}B)^\dagger\|/\tau$ .

**Proof.** Define  $F_K = (I + KP)^{-1}K$  and factorize  $B$  as  $B = GH^T$ , where  $G$  and  $H$  have full column rank. Then

$$\begin{aligned} F_{\hat{K}}(s) &= k (s + kB^T A^{-2} B)^{-1} B^T A^{-2} (sI - A) \\ &= kH (s + kG^T A^{-2} GH^T)^{-1} G^T A^{-2} (sI - A) \end{aligned}$$

so the poles of  $F_K$  are the eigenvalues of  $-kG^T A^{-2} GH^T H$ , which are equal to the non-zero eigenvalues of  $kB^T A^{-2} B$ . In particular,  $F_{\hat{K}}$  and  $P$  are stable, so  $\hat{K}$  is stabilizing.

We also have

$$\begin{aligned} I &\preceq \tau^2 k^2 B^T A^{-2} B \\ B^T (\omega^2 I + A^2)^{-1} B &\preceq \tau^2 [\omega^2 I + k^2 (B^T A^{-2} B)^2] \end{aligned}$$

so

$$\begin{aligned} \tau &\geq \|(i\omega I - A)^{-1} B (i\omega I + kB^T A^{-2} B)^{-1}\| \\ &= \left\| \frac{1}{i\omega} P(i\omega) (I + \hat{K}(i\omega) P(i\omega))^{-1} \right\|. \end{aligned}$$

In general

$$PF_K P = P - P(I + KP)^{-1}$$

so the constraint  $\|\frac{1}{s}P(I + KP)^{-1}\|_\infty \leq \tau$  gives

$$P(0)F_K(0)P(0) = P(0).$$

Hence consider the minimization problem at  $s = 0$ , i.e. to minimize  $\|F\|$  subject to  $P(0)FP(0) = P(0)$ . Standard calculations (Lemma 2 in the appendix) gives that the minimal value  $\gamma = \|P(0)^\dagger\|$  is attained by  $F = P(0)^\dagger$ . In particular, since  $F_{\hat{K}}(0) = P(0)^\dagger$ , it follows that  $\hat{K}$  is optimal at  $\omega = 0$ .

The inequality  $\|F_{\hat{K}}(i\omega)\| \leq \|F_{\hat{K}}(0)\|$  can be rewritten as

$$\begin{aligned} F_{\hat{K}}(i\omega)^* F_{\hat{K}}(i\omega) &\preceq \gamma^2 I \\ A^{-2} B [\omega^2 k^{-2} I + (B^T A^{-2} B)^2]^{-1} B^T A^{-2} &\preceq \gamma^2 (\omega^2 I + A^2)^{-1} \\ k^2 B^T A^{-2} (\omega^2 I + A^2) A^{-2} B &\preceq \gamma^2 [\omega^2 I + k^2 (B^T A^{-2} B)^2] \end{aligned}$$

The last inequality holds trivially for  $\omega = 0$  and it holds for all other  $\omega$  provided that  $k^2 B^T A^{-4} B \preceq \gamma^2 I$ , which is equivalent to the assumption on  $\tau$ . Thus  $\|F_{\hat{K}}\|_\infty$  takes the minimal value  $\gamma$  and the proof is complete.

## 4. CONTROL ON NETWORKS

Theorem 1 can be applied to the following problem:

Given a graph  $(\mathcal{V}, \mathcal{E})$ , suppose that

$$\begin{cases} \dot{x}_i = a_i x_i + b_i u_i + \sum_{(i,j) \in \mathcal{E}} (u_{ij} + v_{ji}) \\ e_i = r_i - x_i \end{cases} \quad (3)$$

$i \in \mathcal{V}$ ,  $a_i < 0$ ,  $x(0) = 0$  and  $u_{ij} = -u_{ji}$ ,  $v_{ij} = -v_{ji}$ . The signals  $u$ ,  $v$ ,  $r$  and  $e$  can be viewed as input-, disturbance-, reference- and error-signals respectively. The problem is to find a control law  $u = k * e$  that minimizes the  $L_2$ -gain from  $r$  to  $u$  while keeping the  $L_2$ -gain from  $v$  to  $z$  bounded by  $\tau$  when  $\dot{z} = x$ ,  $z(0) = 0$ .

The optimal controller  $\hat{K}$  has a distributed realization with one integrator at every node:

$$\begin{cases} \dot{z}_i = e_i \\ u_{ij} = z_i/a_i - e_i/a_i^2 - z_j/a_j + e_j/a_j^2 \\ u_i = b_i(z_i/a_i - e_i/a_i^2) \end{cases} \quad (4)$$

Just like (2), this control law is decentralized in the sense that control action on the edge  $(i, j)$  is entirely determined by the errors at nodes  $i$  and  $j$ . The closed loop map from  $r$  to  $z$  has transfer matrix

$$(sI - A)^{-1} (sI + kBB^T A^{-2})^{-1},$$

and non-negative impulse response. It should be noted that unless the controller realization is minimal, the controller will have integrators that are not stabilized in closed loop. For this reason, it necessary that  $b_i \neq 0$  for at least one node in every connected component of the graph.

## 5. EXAMPLE

Consider the system depicted in Figure 1. The dynamics of the levels in the two buffers marked 1 and 2, around some steady state, is given by

$$\begin{aligned} \dot{x}_1 &= -x_1 + u_1 - u_{12} & e_1 &= r_1 - x_1 \\ \dot{x}_2 &= -2x_2 + u_{12} & e_2 &= r_2 - x_2, \end{aligned} \quad (5)$$

where  $x_1$  is the level in buffer 1 and  $x_2$  is the level in buffer 2. The transfer function of (5) is given by  $P(s) = (sI - A)^{-1}B$  with

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} \quad B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}. \quad (6)$$

The criterion on  $\tau$  is given by  $\tau \geq 1.43$  and  $\hat{K}$  is given by

$$\begin{aligned} u_1(t) &= -ke_1 - k \int_0^t e_1(\sigma) d\sigma \\ u_{12}(t) &= ke_1 - \frac{k}{2} e_2 + k \int_0^t \left( e_1(\sigma) - \frac{1}{4} e_2(\sigma) \right) d\sigma. \end{aligned}$$

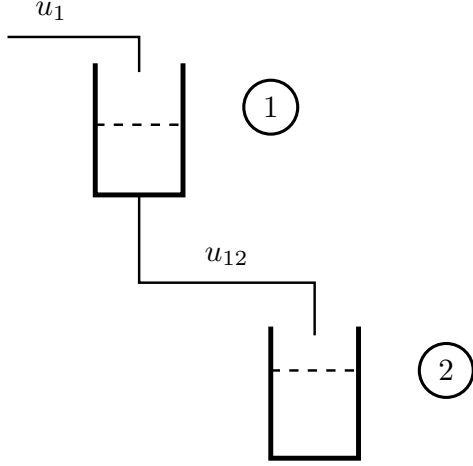


Fig. 1. System with two buffers 1 and 2. The interconnection given by  $B$  in (6) is depicted by the line drawn between the buffers as well as between buffer 1 and the exterior. Degradation due to internal dynamics is not shown.

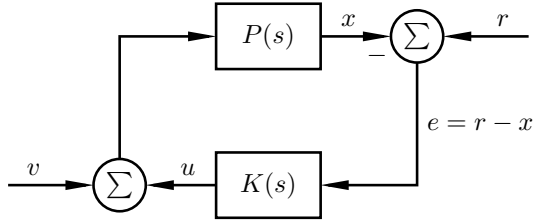


Fig. 2. Block diagram of closed-loop system with measurement  $y$ , control input  $u$ , process disturbance  $v$ , reference value (or measurement error)  $r$ , controller  $K$  and process  $P$ .

for  $k = 8.5/\tau$ . Notice that each control input is only using the state(s) it affects through the matrix  $B$ , i.e., the controller has the same zero-block structure as  $B^T$ . Thus, the controller only considers local information, where local is subject to the interconnection specified by  $B$ .

The optimization objective stated in Theorem 1 concerns the transfer function

$$(I + KP)^{-1}K,$$

which maps the reference value  $w$  to the control input  $u$ , as depicted in Figure 2. Thus, the first objective is to minimize the control effort needed to follow the reference value  $w$ . The second performance criterion

$$\left\| \frac{1}{s} P(I + KP)^{-1} \right\|_{\infty} \leq \tau \quad (7)$$

specifies the control quality in terms of disturbance rejection. The impact from a low frequency process disturbance  $v$  should be attenuated by the feedback loop. The parameter  $\tau$  is a time constant that determines the bandwidth of the control loop. The impact of  $\tau$  will be illustrated below.

Given

$$P(s) = (sI - A)^{-1}B$$

$$\hat{K}(s) = \frac{\|(A^{-1}B)^{\dagger}\|}{\tau} \left( B^T A^{-2} - \frac{1}{s} B^T A^{-1} \right),$$

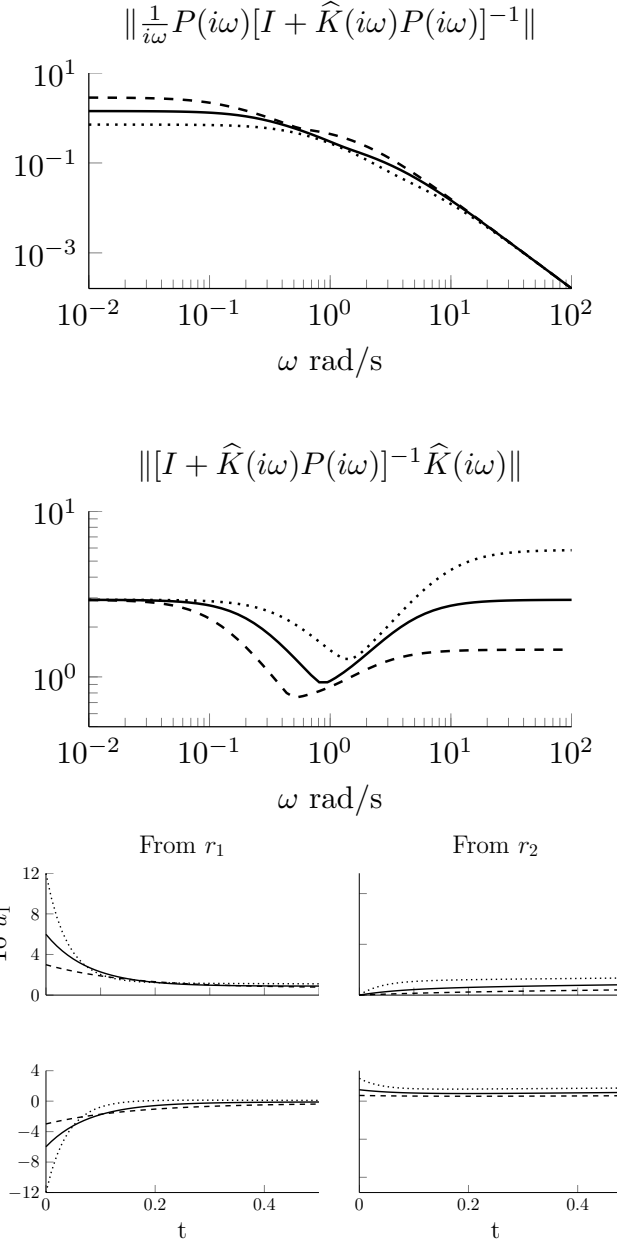


Fig. 3. Spectral norms and step responses for different  $\tau$ ; define  $\tau_* = \sqrt{\|B^T A^{-4} B\|}$ , then  $\tau = \tau_*$  gives the solid line,  $\tau = 2\tau_*$  the dashed line and  $\tau = \tau_*/2$  the dotted line. The last case violates the optimality conditions of Theorem 1 and the norm is not maximal at  $\omega = 0$ .

Figure 3 plots the norms of  $\|\frac{1}{i\omega}P(i\omega)[I + \hat{K}(i\omega)P(i\omega)]^{-1}\|$  and  $\|[I + \hat{K}(i\omega)P(i\omega)]^{-1}\hat{K}(i\omega)\|$  against the frequency  $\omega$  for three different values of  $\tau$ .

The first diagram clearly shows that the disturbance rejection is increasingly effective as the time constant  $\tau$  is reduced. However, as shown in the second diagram, this comes at the price of larger control signals at higher frequencies, while the gain at  $\omega = 0$  remains unchanged. In the step response diagrams, a smaller value of  $\tau$  results in larger control values for small  $t$ , while the steady state values remain unchanged.

## 6. CONCLUSIONS

This paper has formulated a class of dynamic state feedback control problems for which a structure preserving PI controller is  $H_\infty$  optimal. An explicit expression for the optimal gain has been given, which clarifies the relationship between plant structure and achievable performance.

## ACKNOWLEDGEMENTS

This work was supported by the Swedish Research Council through the LCCC Linnaeus Center. The authors are members of the LCCC Linnaeus Center and the eLLIIT Excellence Center at Lund University.

## REFERENCES

- Bassam Bamieh, Fernando Paganini, and Munther A. Dahleh. Distributed control of spatially invariant systems. *IEEE Transactions on Automatic Control*, 47(7), July 2002.
- Corentin Briat. Robust stability and stabilization of uncertain linear positive systems via integral linear constraints:  $L_1$ -gain and  $L_\infty$ -gain characterization. *International Journal of Robust and Nonlinear Control*, 23: 1932–1954, 2013.
- Raffaello D’Andrea and Geir E Dullerud. Distributed control design for spatially interconnected systems. *IEEE Transactions on automatic control*, 48(9):1478–1495, 2003.
- Gunther Dirr, Hiroshi Ito, Anders Rantzer, and Björn Rüffer. Separable Lyapunov functions for monotone systems: Constructions and limitations. *Discrete and Continuous Dynamical Systems. Series B*, 20(8), 2015.
- Carolina Lidström and Anders Rantzer. Optimal  $H_\infty$  state feedback for systems with symmetric and Hurwitz state matrix. In *Proceedings of American Control Conference*, Boston, July 2016.
- Anders Rantzer. Scalable control of positive systems. *European Journal of Control*, 24:72–80, 2015.
- Michael Rotkowitz and Sanjay Lall. Decentralized control information structures preserved under feedback. In *Proceedings of the IEEE Conference on Decision and Control*, December 2002.
- Takashi Tanaka and Cédric Langbort. The bounded real lemma for internally positive systems and  $H_\infty$  structured static state feedback. *IEEE Transactions on Automatic Control*, 56(9):2218–2223, September 2011.

## 7. APPENDIX

*Lemma 2.* Let  $A \in \mathbb{C}^{n \times m}$ . Then

$$\min_{X \in \mathbb{C}^{m \times n}} \|X\| \quad \text{s.t.} \quad AXA = A \quad (8)$$

has the minimal value  $\|A^\dagger\|$ , attained by  $\hat{X} = A^\dagger$ .

**Proof.** Let  $y \in \text{Im}(A)$  be a unit vector in  $\mathbb{C}^m$  and let  $X$  be any feasible point for (8). Consider:

$$\min_{x \in \mathbb{C}^n} |x| \quad \text{s.t.} \quad y = Ax \text{ and } x = Xy. \quad (9)$$

Observe that because  $X$  is feasible, the optimal solution always exists. The value equals  $|Xy|$ , giving a lower bound for  $\|X\|$ . Relaxing (9) by removing the second constraint gives a least squares problem with solution  $x = A^\dagger y$  and

value  $|A^\dagger y|$ . Maximizing over  $y$  gives  $\|A^\dagger\|$ . The result follows because  $X = A^\dagger$  achieves this lower bound.